Deep Learning based Zero Watermarking for Authentication of Medical Records

Gurleen Kaur, Bakul Gupta, and Ashima Anand

Computer Science and Engineering, Thapar Institute of Engineering and Technology

Abstract. The security of digital images is crucial since they often contain sensitive and confidential data. Unauthorized access to this data could result in severe penalties for the parties involved. Despite the availability of highly secure algorithms, security remains a significant concern due to the rapid emergence of new technologies that can breach it. Thus the proposed work implements a technique that makes the confidential data inaccessible to intruders. Hence fragile type of data hiding technique is used where even with the slightest tampering to the image by an attacker, the information i.e. watermark image is completely destroyed, hence preventing it from unauthorized access. Also, a hybrid transform including DTCWT and NSST is used to fuse two medical images to form a more sophisticated output image, which serves as the final watermark. Further, the zero watermarking model is implemented using the ResNet 50 DL model for more precise results and extraction of feature maps. Embedding the actual image in the carrier image could make the watermarking detectable especially when it is fragile, hence Zero Watermarking overcomes this also by virtual embedding. Moreover, the algorithm employs the avalanche effect of SHA512 for highly secure authentication, further strengthening the security of the system. Overall, the proposed method is an effective way to ensure the security of digital images with confidential data.

Keywords: Zero watermarking, Image Fusion, RDWT, Encryption, Medical images, Deep Learning.

1 Introduction

In today's digital age, the use of digital images has become ubiquitous in every sector of society, ranging from personal photography to medical imaging, from social media to e-commerce. With the increasing use of digital images, the need for their security has also become paramount. Digital images contain sensitive and confidential information, which if compromised, can lead to significant consequences such as identity theft, loss of personal privacy, and even financial losses. Therefore, it is imperative to ensure that digital images are adequately protected from unauthorized access, manipulation, and theft.This paper emphasizes the critical importance of ensuring the security of medical images, highlighting their vulnerability to unauthorized access and potential misuse. Various methods are employed for medical diagnosis, including ultrasonography, magnetic resonance imaging, positron emission tomography etc. Diagnostic images undergo an extensive array of processes, encompassing tasks such as feature selection, image denoising, and segmentation, and they are extensively archived and distributed [1].

One method of protecting digital images is through the use of watermarking. Various conventional techniques of watermarking have been employed by researchers to safeguard copyright in domains that are both fragile and robust.[2]. These techniques vary in their intended function and the level of security that they afford to digital data, as depicted in Figure 1.

Fragile watermarking is most effective in situations where digital media authentication is imperative. It employs a watermark as a digital signature, thereby validating the authenticity of the media and ensuring that it has not been altered. This feature makes it highly valuable in contexts of data authentication, where it is necessary to verify the genuineness of a document or image. [3]. On the contrary, robust watermarking is formulated to endure

David C. Wyld et al. (Eds): ICCSEA, NLAI, SCAI, CSIA, IBCOM, SEMIT, NECO, SPPR, MLDS -2023
pp. 129 - 141, 2023.CS & IT-CSCP 2023
DOI: 10.5121/csit.2023.132411

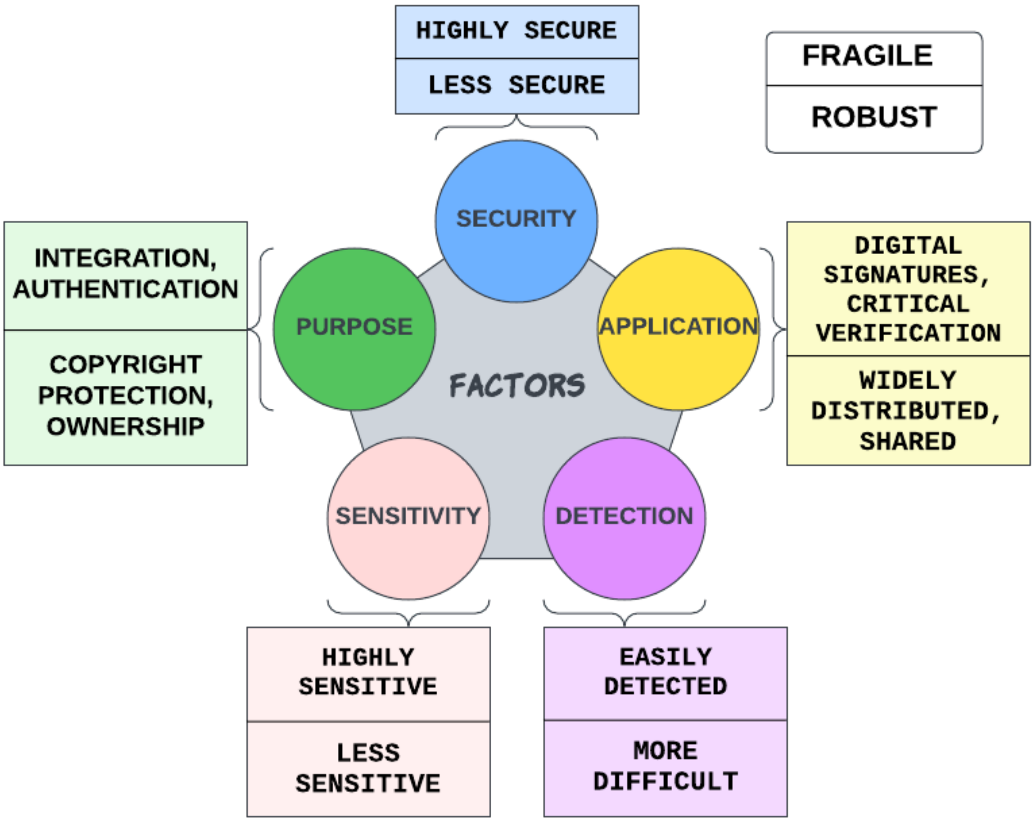


Fig. 1. Difference between Fragile and Robust Watermarking

typical signal processing attacks like compression, cropping, and filtering so that digital data still can be extracted. This makes it ideal for applications where the digital media needs to be distributed and shared widely, such as in the entertainment industry. However, the robustness of the watermark comes at the cost of reduced sensitivity, which means that it may not be able to detect small changes to the original digital media. Whereas, fragile watermarking is done in such a way that watermark data is completely destroyed after it encounters any change, it is harder for an attacker to modify or remove them without detection. Therefore, Fragile watermarking offers a higher level of security than robust watermarking. Therefore, this paper has opted for the fragile watermarking technique as the chosen approach for the necessary algorithm.

Over the recent years, researchers have extensively employed Zero watermarking, which can be thought of as "invisible watermarking." The usage of the term "zero" implies that the embedded watermark is crafted to remain unseen or imperceptible to human senses, such as sight or hearing, without causing any significant change in the original content's visual or auditory attributes. However, a significant portion of the research has centered around utilizing Zero watermarking as the robust algorithm. Hence, the authors undertook this study to explore the outcomes when the imperceptibility of Zero watermarking is integrated with fragile watermarking. The primary objective was to address the drawback of fragile watermarking, which is its susceptibility to detection compared to robust watermarking techniques.

In real-world scenarios, physicians often require multiple patient reports to make an accurate diagnosis. Taking this aspect into account, the present study has opted to consider that patients need to submit both their CT scan and MRI images. Addressing this concern, the paper introduces a fusion algorithm that can be described as a form of encryption and also a multi-factor authentication method.

The objective of this manuscript is to present a technique for securely transmitting digital images in a manner that ensures confidential information remains inaccessible to intruders and maintainting the authenticity of images.

2 Literature Survey

In recent years, with remarkable progress in deep learning techniques, they have also been extensively employed in safeguarding digital information.[4] paper proposed the first watermarking framework using CNN.It introduces a novel non-blind digital image watermarking method that utilizes the auto-encoder capability of CNNs. This approach generates positive and negative codebook images, which play a crucial role in embedding and extracting watermarks. But as their proposed method was completely non-blind, it lacks practicality in real world. DFT based Zero watermarking along with VGG19 and perceptual hashing was presented by authors in paper [5]. Though their framework was robust against various geometric attacks, but a CNN residual network with more deeper layers could give more accurate results than VGG19 which has 19 layers. Hence, within the algorithm presented in our paper, we opt for the utilization of Residual networks.

Authors of Ref. [6] proposed the method for Zero Watermarking with DCT and Residual DenseNet. Their proposed framework was robust against various geometric attacks. Further DWT-SVD-DCT based fragile watermarking was implemented in paper [7]. The authors developed a tamper-proof framework that could identify cropping and object insertion attacks. To enhance the framework's sensitivity and detect even minor alterations, an authentication code was generated and incorporated into the watermarked image using the QIM technique. The embedded code was extracted using the Gram-Schmidt process. However, it should be noted that the Gram-Schmidt process may result in information loss, which could erroneously detect an attack even if it did not occur. Further the paper has not analyzed results for many signal processing attacks.

Singh et al. [8] has put forward fragile watermarking technique using DCT-LSB method. While the paper's method is capable of accurately extracting the main content from a tampered image up to a 50% tampering rate, it is only effective in restoring against certain types of attacks. It cannot fully remove the watermark in cases of tampering, which is a crucial requirement for fragile watermarking as it must be highly responsive. Hence for making the algorithm highly sensitive and sensitive, hashing could also play the major role in authentiation. [9] paper completed the comparatively analysis of well known MD and SHA algorithms. Further concluded that for high security purpose SHA512 could serve the purpose because of its avalanche effect and longer length of string constructed.

Summing up the comprehensive review of existing literature, we deduced that Zero watermarking has predominantly been applied in robust methods. However, a drawback lies in the fact that while robust watermarking offers enhanced security, fragile watermarking surpasses it in terms of security measures. When it comes to employing watermarking techniques rooted in Deep Learning, a constant compromise between time complexity and result accuracy is unavoidable. As a result, our initial focus rests on the evaluation of outcomes across different models, as detailed in subsequent sections. To bolster security, the integration of the SHA hashing algorithm stands out as a promising choice.

3 Proposed Methodology

This paper puts forth a fragile watermarking technique that is well-suited for binary images of a confidential nature (See Fig. 2). Until now, the majority of the existing fragile

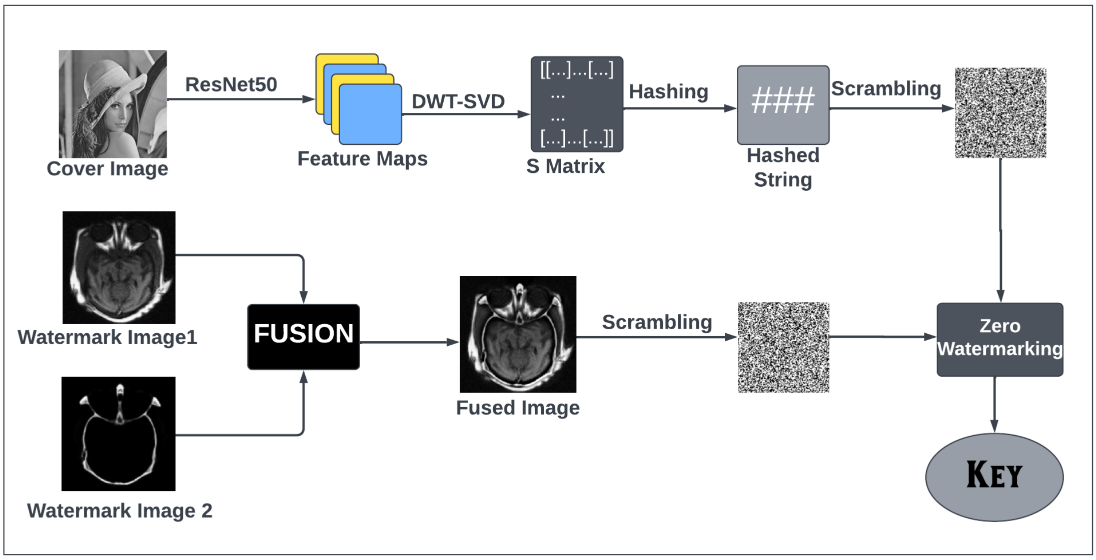


Fig. 2. Framework of the proposed fragile watermarking model using image fusion

watermarking methods that embedded the watermark were easily detectable. However, the proposed approach employs zero watermarking, which doesn't embed the watermark directly but instead relies on the resemblances between the cover image and the watermark image. Unlike conventional zero watermarking methods, the proposed approach abstains from manual feature extraction and instead leverages the power of deep learning in conjunction with zero watermarking. We have devised a methodology that entails a 3 step process. Initially, fusion of two distinct watermarks is executed, resulting in the creation of a solitary watermark. Subsequently, the extraction of a feature map from the cover image is carried out. Finally, the zero watermark technique is implemented to accomplish the watermarking process. For the purpose of further enhancing security,hashing and scrambling of images is also carried out.

3.1 Watermark Generation using NSST-DTCWT transforms based Image Fusion

In this phase, DTCWT and NSST-based fusion of medical images is implemented to generate the watermark using two input images, 'CT' and 'MRI'. This image fusion algorithm uses two different rule sets to fuse the high and low-frequency coefficients of the input images. Parameter adaptive PCNN is used to fuse the high-frequency coefficients, while WSE and WSNML-based rules contribute to generating the fused low-frequency coefficients. The flowchart for generating the fused image as a mark carrier is shown in Fig.3.

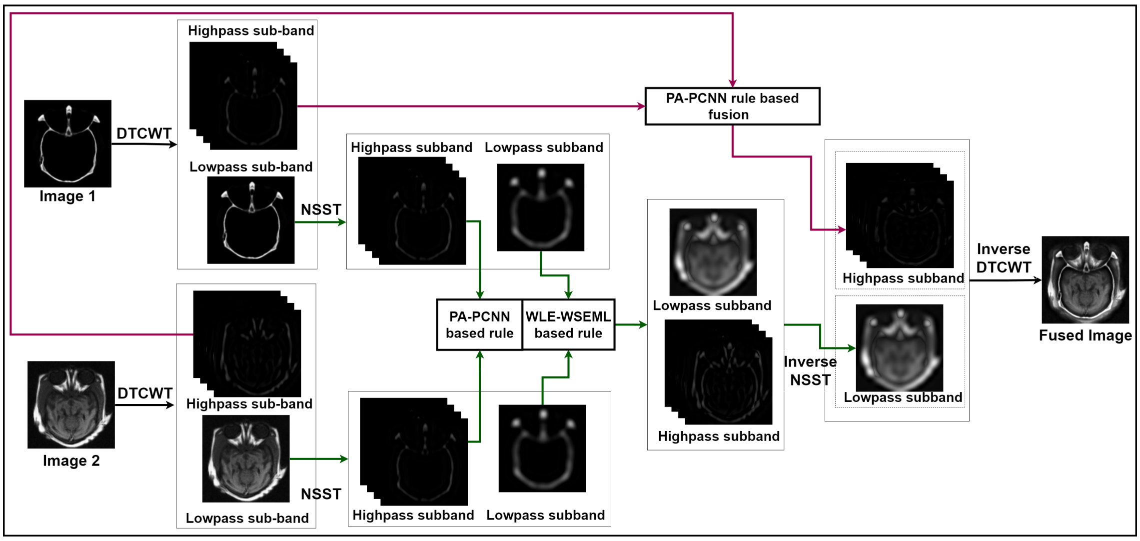


Fig. 3. DTCWT-NSST based Image Fusion to generate final Watermark Image

To fuse the high-frequency coefficients, PA-PCNN-based fusion rules are adapted. PA-PCNN is a type of PCNN that works by monitoring the image processing performance and dynamically adjusting the parameters to optimize the results. This adaptation is performed through a feedback mechanism that adjusts the parameters based on the current state of the image processing. The parameters that are adapted in PA-PCNN include the coupling strength, the threshold and the pulse width.

PA-PCNNs can be used to preserve the relevant information from each image while suppressing the irrelevant information by automatically adjusting their parameters based on the input data. This enables the network to extract and preserve the most relevant features from each image, resulting in a fused image that contains more information than any of the individual images.

Further, the edge and contours of the input images are preserved by fusing the low-frequency coefficients of input images using a hybrid of WLE and WSEML. The activity level measure, WLE, is mathematically calculated as,

$$\mathrm{WLE}_{ab}(x,y) = \sum_{i}\sum_{j} Wm'(i,j) H_{ab} \times (x+i, y+j)^2, \tag{1}$$

where $H_{ab}(x,y)$ denotes the high-frequency NSST coefficient at position *(x, y)* of direction b at layer a. Also, the Wm' is the weighted matrix which is defined as:

$$Wm' = \frac{1}{16}\begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix} \tag{2}$$

Further, WSEML is used to extract the details of the input images using the following equation,

$$\mathrm{WSEML}_{a,b}(x,y) = \sum_{p=-rad}^{rad}\sum_{q=-rad}^{rad} (Wm'(p+rad+1, q+rad+1) \times \mathrm{EML}_{a,b}(a+p, b+q)) \tag{3}$$

where, Wm' is the weighting matrix defined in eq. 2, and EML is defined as,

$$\begin{aligned}\mathrm{EML}_{img}(x,y) = |2img(x,y) - img(x-1,y) - img(x+1,y)| \\ + |2img(x,y) - img(x,y-1) - img(x,y+1)| \\ + \frac{1}{\sqrt{2}}|2img(x,y) - img(x-1,y-1) - img(x+1,y+1)| \\ + \frac{1}{\sqrt{2}}|2img(x,y) - img(x-1,y+1) - img(x+1,y-1)| \quad (4)\end{aligned}$$

Table 1. Details of notations used in this article

Notation	Explanation	Notation	Explanation
CT, MRI	Input images for multimodality image fusion	dtcwt_L1, dtcwt_H1	High and Low DTCWT coefficients of CT image
nsst_L1, nsst_H1	Low and high NSST coefficients of dtcwt_1	WLE1, WSEML1	WLE and WSEML associated with dtcwt_H1 and dtcwt_H2
nsst_L	Fused low frequency NSST coefficient	nsst_H	Fused high frequency NSST coefficient
dtcwt_H	Fused high frequency DTCWT coefficient	dtcwt_L	Fused low frequency DTCWT coefficient
Fimg	Fused Image	cI	Cover Image
key	Extraction Key	fusedWI	fused Watermark Image after resizing
s_WI	Watermark Image after applying seed scrambling	ResNet50	Residual Network with 50 layers
model	contains weights of ResNet50 pre-trained on imagenet dataset	FM	Feature Matrix of the cover image predicted using 'model'
DWT	Discrete Wavelet Transformation	SVD	Singular Value Decomposition
cA, cH, cV, cD	Approximation, Vertical, Horizontal and Diagonal sub-bands of 'FM' on applying DWT	u,s,v	Left, Middle and Right singular matrices of 'cA' on applying 'SVD'
b_cover	Binary of the cover image after hashing	bin_wat	Binary of the watermark image
cov	Binary of cover Image after adjusting its length	s_cov	Cover Image after applying seed scrambling

As shown in Algorithm 1, two input images, 'CT' and 'MRI', are decomposed into high-pass and low-pass components using DTCWT transform. The resultant high-pass components, 'dtcwt_H1' and 'dtcwt_H2', are merged using a parameter adaptive PCNN (PAPCNN)-based fusion scheme, resulting in 'dtcwt_H'. Further, NSST is applied to the low-pass DTCWT coefficients. The resultant high-band NSST coefficients are again fused based on fusion rules using PAPCNN. The energy preserving and detail extracting issues are addressed by fusing the low-band NSST coefficients using WLE and WSEML-based fusion rules, generating the fused low-band NSST component, 'nsst_L'. Inverse NSST is then applied to form the fused low-pass DTCWT coefficients, 'dtcwt_L'. Finally, inverse DTCWT is applied to obtain the fused image, 'Fimg', which is treated as watermark image in the later sections.

3.2 Extraction of Feature maps

The extraction of feature maps is a crucial process that necessitates meticulous attention and consideration to guarantee the precision, dependability, and utility of the resulting

Algorithm 1: Algorithm of DTCWT-NSST based medical image fusion

```
Input  : CT, MRI
Output: Fimg
// Phase 1: Transforming input images using DTCWT
   dtcwt_L1, dtcwt_H1 ← DTCWT(CT);
   dtcwt_L2, dtcwt_H2 ← DTCWT(MRI);
// Phase 2: Transforming low-frequency sub-band using NSST
   nsst_L1, nsst_H1 ← NSST(dtcwt_L1);
   nsst_L2, nsst_H2 ← NSST(dtcwt_L2);
// Phase 3: Fusion of low-frequency NSST coefficients
   WLE1 ← WLE_Calculation(nsst_L1);
   WLE2 ← WLE_Calculation(nsst_L2);
   WSEML1 ← WSEML_Calculation(nsst_L1);
   WSEML2 ← WSEML_Calculation(nsst_L2);
   map ← (WLE1 × WSEML1 ≥ WLE ∗ WSEML2);
   nsst_L ← map. ∗ nsst_L1 +  map. ∗ nsst_L2;
// Phase 4: Fusion of high-frequency NSST coefficients
   nsst_P1 ← PA_PCNN(nsst_H1);
   nsst_P2 ← PA_PCNN(nsst_H2);
   map ← (nsst_P1 >= nsst_P2);
   nsst_H ← map. ∗ nsst_P1 +  map. ∗ nsst_P2;
// Phase 5: Applying inverse NSST decomposition
   dtcwt_L ← Inverse_NSST(nsst_L, nsst_H);
// Phase 6: Fusion of high-frequency DTCWT coefficients
   dtcwt_P1 ← PA_PCNN(dtcwt_H1);
   dtcwt_P2 ← PA_PCNN(dtcwt_H2);
   map ← (dtcwt_P1 >= dtcwt_P2);
   dtcwt_H ← map. ∗ dtcwt_P1 +  map. ∗ dtcwt_P2;
// Phase 7: Applying inverse DTCWT decomposition
   Fimg ← Inverse_DTCWT(dtcwt_L, dtcwt_H);
return Fimg
```

features in required applications.

The approach taken in this paper involves the utilization of ResNet50 whose building block is shown in fig.4 , a convolutional neural network (CNN) with a significant depth of 50 layers. The depth of the network is crucial for neural networks, but deeper networks are

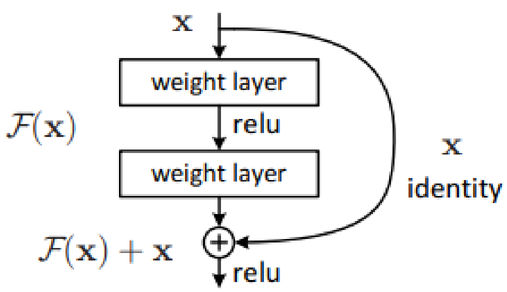


Fig. 4. Residual learning: a building block[10]

more difficult to train. The configuration of ResNet50 enables the instruction of networks and permits them to be considerably more profound, resulting in augmented proficiency in various tasks. ResNet50 surpasses their simple equivalents in depth, and furthermore, the quantity of weights in such networks is significantly lower. [11].

When a 512x512 cover Image *'cI'* is fed through a modified ResNet50 model, the last convolutional layer produces a feature map with a spatial size of 16x16 and 2048 channels. Each channel in the feature map represents a specific learned feature of the input image. To obtain a final feature matrix, the individual feature maps are added element-wise to produce a single feature map. This final feature map preserves the spatial structure of the input image and contains information about the learned features across all channels. This feature map is used for further processing.

In ResNet50, the feature maps are extracted through a series of convolutional layers that are arranged in blocks. Each block consists of multiple layers, including convolutional layers, batch normalization layers, and activation layers. The output of each block is then passed through a shortcut connection that allows the gradient to flow more easily during training. In proposed method, we have used pre-trained model of ResNet50 on Imagenet Dataset.

3.3 Zero Watermarking

As described in[12], Zero watermarking essentially performs a virtual embedding process, where the key for generating a watermark is generated by analyzing the similarities between the attributes of the cover image and the watermark image. This key can then be used to produce an identical watermark from the cover image.

Method used for Zero Watermarking is described in Algorithm 2. The fused image '*Fimg*'

Algorithm 2: Watermarking

Input : Fimg, cI
Output: key

```
// Phase 1: Resizing of input images
    fusedWI ← resize(Fimg, (256, 256));
    cI ← resize(cI, (512, 512));
// Phase 2: Scrambling of fusedWI
    s_WI ← seed_scramble(fusedWI);
// Phase 3: Feature Map Extraction
    model ← ResNet50(weights = 'magenet');
    cI ← resnet50.preprocess_input(cI);
    FM ← model.predict(cI);
// Phase 4: Apply DWT and SVD
    [cA, cH, cV, cD] ← dwt(FM) ;
    [u, s, v] ← svd(cA);
// Phase 5: Hashing Cover Image
    b_cover ← SHA512(s);
// Phase 6: Adjusting length of binary of cover image
    bin_wat ← matrix_to_binary(s_WI);
    cov ← add_trailing_zeros(b_cover, bin_wat);
// Phase 7: Scrambling binary of cover image
    s_cov ← seed_scramble(cov);
// Phase 8: Key Generation
    key ← XOR(s_cov, bin_wat);
return key
```

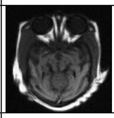
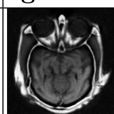
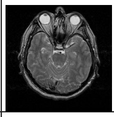
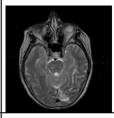
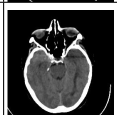
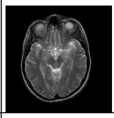
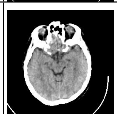
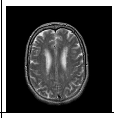
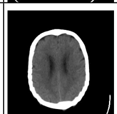
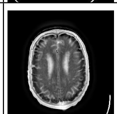
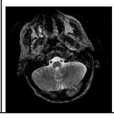
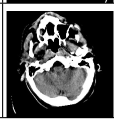
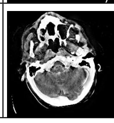
obtained after combining 'CT' and 'MRI' in Algorithm 1 serves as the watermark image. It is used as input along with the cover image 'cI'. After resizing, '*Fimg*' is scrambled using a seed scrambling algorithm. A feature map is extracted from '*cI*' using the ResNet50 DL model that has been pre-trained on the imagenet dataset. This feature map '*FM*' is then divided into frequency subbands with DWT, and the '*LL*' subband is subject to SVD decomposition. The resulting '*s*' singular value matrix is hashed with the SHA512 algorithm. Both the scrambled watermark image '*s_WI*' and the hashed string '*b_cover*' are converted to binary form. To account for differences in size between the two binary outputs, a string of trailing zeroes is added to '*b_cover*', and the same scrambling algorithm as previously mentioned is applied. The XOR operation is then performed on '*bin_wat*' and '*s_cov*' to generate the extraction key '*key*'.

4 Results and Analysis

This part of the analysis thoroughly analyzes the suggested technique, beginning with experimental configurations and progressing to performance evaluation with varied cover images and fusion images. The comparative study of the proposed technique is offered at the end.

The trial begins initially by taking a brain MRI picture sized 512×512 as the cover object and the covert data is taken as a CT Scan image sized 256×256. The implementation is done on MATLAB R2021b using system with the following configuration, Intel Xeon(R) Gold processor with 256GB RAM. The performance metrics used for assessing the projected technique are listed as (PSNR), SSIM, and NC. The evaluation parameters of Fusion method quantifies how accurately the fused image conveys the original images content. Some common fusion parameters are MI, QABF, SSIM, SF, and STD [13–15]. Visual output for 6 sample input output pairs are shown in table 2.

Table 2. Sample of images used for evaluating Dtcwt-Nsst based multi-modality image fusion

Pair No.	**Image 1**	**Image 2**	**Fused Image**
Pair 1			
Pair 2			
Pair 3			
Pair 4			
Pair 5			
Pair 6			

The objective evaluation of the proposed DTCWT-NSST based fusion method, when implemented on 50 pairs of medical images [16], is referred to in Table 3. Average scores

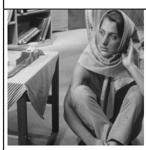
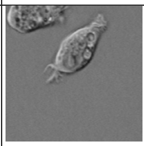
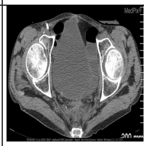
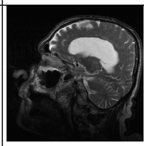
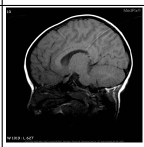
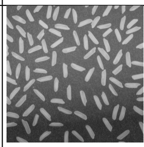
Table 3. Evaluation of proposed image fusion method on 50 pair of medical images

Pair	Entropy	MI	QABF	FMI	NABF	SSIM	SF	STD	PSNR1	PSNR2
Pair 1	1.7126	3.4252	0.1937	0.8895	0.0373	0.0932	4.2698	36.2066	59.8872	66.1457
Pair 2	6.7562	13.5125	0.4596	0.8789	0.2290	0.4766	6.0882	54.8540	70.0475	59.1952
Pair 3	2.9538	5.9075	0.3288	0.8700	0.0452	0.6626	5.4185	84.7166	65.8571	59.2004
Pair 4	4.4879	8.9757	0.2748	0.8929	0.0731	0.8234	4.5951	57.0351	70.13	65.7703
Pair 5	4.8135	9.6269	0.3957	0.8821	0.0769	0.6659	5.8501	83.0159	65.8074	59.3964
Pair 6	4.3270	8.6540	0.3049	0.8672	0.0666	0.5857	6.5350	80.7836	65.4172	58.8764
Average of 50 pairs	3.8406	7.6811	0.2906	0.8771	0.0419	0.6653	4.9579	79.5419	67.1835	59.9599

of QABF, FMI, SSIM, SF, and STD for 50 pairs are 0.2906, 0.8771, 0.6653, 4.9579 and 79.5419, respectively. These results verify the satisfactory performance of the multi-modality fusion method for healthcare images.

Table 4 displays a range of standard images employed for analyzing outcomes, while a medical image data-set of over 250 MRI images is also utilized for the same objective.

Table 4. Standard Images

I1	I3	I5	I7	I9	I11
					
I2	**I4**	**I6**	**I8**	**I10**	**I12**
					Average of 250+ MRI Images

The proposed method is being tested against various geometric and signal processing attacks, as illustrated in table 5. NC value is calculated original watermark and extracted

Table 5. Attacks performed

Symbol	Attack Names
A1	Average filtering
A2	Cropping Attacks
A3	Gaussian filter
A4	JPEG compression
A5	Median Filter Attack
A6	Rotation Attack
A7	Salt & Pepper Noise
A8	Scaling the image
A9	Translation attack

watermark which is the measure of similarity between images. In the case of fragile watermarking, the NC value between the watermark and extracted image should be very low even after minimal modification to the watermarked image, serving the purpose of privacy and security. As the proposed method employs zero watermarking, there is no actual embedding in the cover image, and thus no metric is required to compare the watermarked and cover images as they are the same.

In order to extract feature maps from the cover image, diverse deep learning (DL) models were computed and executed on standard images in the face of attacks. These outcomes were then meticulously analyzed, and the average results are elucidated in table 6.

Table 6. Average NC after performing attacks on various DL architectures

DL Tech.	A1	A2	A3	A4	A5	A6	A7	A8	A9
R50	0.0159	-0.0175	0.0078	-0.0054	-0.0275	-0.0435	0.0103	-0.0068	-0.0005
R101	-0.0146	0.0008	-0.0088	0.0056	-0.0010	-0.0039	-0.0021	-0.0076	-0.0056
VGG19	0.0070	0.0024	-0.0167	-0.0177	0.0037	0.0243	-0.0220	-0.0062	-0.0091
DN121	0.0071	-0.0021	-0.0087	-0.0102	-0.0142	0.0010	0.0036	0.0010	-0.0138
MNet	0.0200	0.0295	-0.0039	-0.0024	-0.0191	-0.0044	-0.0079	-0.0130	0.0090

Subsequent to running ResNet50, ResNet101, VGG19, DenseNet121, and MobileNet models, the deduction drawn from the results was that ResNet50 provided the most exceptional outcomes since it generated the minimum value of NC is majority of attacks and is giving best results from all other methods.

Furthermore, we also determined the average value of NC using ResNet50, by testing it on a medical image dataset against various attacks. The outcomes of ResNet 50 against various attacks on individual standard images and medical image dataset are tabulated in table 7. Hence the NC values obtained from the results are very low, this directly prove that the proposed method is highly tamper proof.

Table 7. Results of ResNet50

Images	A1	A2	A3	A4	A5	A6	A7	A8	A9
I1	0.0893	-0.1022	-0.0673	0.0288	-0.0300	-0.0306	0.0006	-0.0187	0.0307
I2	0.0726	-0.0253	0.0685	-0.0287	0.0131	-0.0823	-0.0091	-0.0118	-0.0486
I3	-0.0527	0.0047	0.0454	0.0252	-0.0053	-0.0465	0.0373	-0.0172	-0.0717
I4	-0.0860	-0.1110	0.0264	-0.0617	-0.0043	-0.0212	-0.0277	-0.0446	0.0010
I5	-0.0370	0.1179	0.0020	-0.0136	0.0108	-0.0347	-0.0195	-0.0602	-0.0190
I6	0.0347	-0.0230	0.0068	-0.0541	-0.0058	-0.0195	0.0786	0.0497	-0.0157
I7	0.0332	-0.0556	-0.0425	-0.0198	-0.0251	-0.0956	0.0512	-0.0076	0.0545
I8	0.0262	-0.0224	-0.0531	0.0029	-0.0726	-0.0383	0.0339	-0.0449	-0.0112
I9	0.0913	-0.0205	0.0154	0.0037	-0.0476	-0.0559	0.0177	0.0131	-0.0158
I10	0.0094	0.0779	0.0037	0.0200	-0.0669	-0.0405	-0.0163	0.0131	0.0269
I11	-0.0060	-0.0325	0.0807	0.0376	-0.0692	-0.0134	-0.0337	0.0544	0.0636
I12	0.00005	-0.00366	0.00492	-0.00118	0.00036	0.00490	-0.00352	-0.00609	0.00124

5 Conclusion

In conclusion, the security of digital images with sensitive and confidential data is of paramount importance to avoid severe penalties associated with unauthorized access. The proposed method in this manuscript is a highly effective solution to this problem. This technique initially employs NSST-DTCWT based multimodality image fusion method to generate the final watermark. By utilizing the avalanche effect of the hashing algorithm SHA512, the method is highly fragile. Any slight changes in the watermarked image can completely destroy the confidential information stored as the watermark image, effectively preventing any unauthorized access. Further, the implementation of virtual embedding with zero watermarking instead of actual embedding significantly reduces the attack surface as it conceals the availability of the watermark image in the cover image. This approach provides an additional layer of security by making it more difficult for attackers to identify and tamper with the watermark image. Therefore, this method provides an excellent solution to ensure the security of digital images with confidential information.

References

1. A. Odeh and Q. A. Al-Haija, "Medical image encryption techniques: a technical survey and potential challenges," *no. January*, pp. 3170–3177, 2023.
2. M. Begum and M. S. Uddin, "Digital image watermarking techniques: a review," *Information*, vol. 11, no. 2, p. 110, 2020.
3. J. Fridrich, "Methods for tamper detection in digital images," in *Multimedia and Security, Workshop at ACM Multimedia*, vol. 99, pp. 29–34, 1999.
4. H. Kandi, D. Mishra, and S. R. S. Gorthi, "Exploring the learning capabilities of convolutional neural networks for robust image watermarking," *Computers & Security*, vol. 65, pp. 247–268, 2017.
5. B. Han, J. Du, Y. Jia, and H. Zhu, "Zero-watermarking algorithm for medical image based on vgg19 deep convolution neural network," *Journal of Healthcare Engineering*, vol. 2021, 2021.
6. C. Gong, J. Liu, M. Gong, J. Li, U. A. Bhatti, and J. Ma, "Robust medical zero-watermarking algorithm based on residual-densenet," *IET Biometrics*, vol. 11, no. 6, pp. 547–556, 2022.
7. T.-S. Nguyen, "Fragile watermarking for image authentication based on dwt-svd-dct techniques," *Multimedia Tools and Applications*, vol. 80, no. 16, pp. 25107–25119, 2021.
8. D. Singh and S. K. Singh, "Dct based efficient fragile watermarking scheme for image authentication and restoration," *Multimedia Tools and Applications*, vol. 76, pp. 953–977, 2017.
9. S. Long, "A comparative analysis of the application of hashing encryption algorithms for md5, sha-1, and sha-512," in *Journal of Physics: Conference Series*, vol. 1314, p. 012210, IOP Publishing, 2019.
10. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
11. Y. Chu, X. Yue, L. Yu, M. Sergei, and Z. Wang, "Automatic image captioning based on resnet50 and lstm with soft attention," *Wireless Communications and Mobile Computing*, vol. 2020, pp. 1–7, 2020.
12. Y. Zhou and W. Jin, "A novel image zero-watermarking scheme based on dwt-svd," in *2011 International Conference on Multimedia Technology*, pp. 2873–2876, IEEE, 2011.
13. N. Jain, A. Yadav, Y. Kumar Sariya, and A. Balodi, "Analysis of discrete wavelet transforms variants for the fusion of ct and mri images," *The Open Biomedical Engineering Journal*, vol. 15, no. 1, 2021.
14. W. Ma, K. Wang, J. Li, S. X. Yang, J. Li, L. Song, and Q. Li, "Infrared and visible image fusion technology and application: A review," *Sensors*, vol. 23, no. 2, p. 599, 2023.
15. L. Tang, Y. Deng, Y. Ma, J. Huang, and J. Ma, "Superfusion: A versatile image registration and fusion network with semantic awareness," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 12, pp. 2121–2137, 2022.
16. J. A. B. Keith A. Johnson, "the whole brain atlas." `http://www.med.harvard.edu/AANLIB/home.html`.

Authors

Gurleen Kaur is currently an undergraduate 4th year Computer Engineering student at Thapar Institute of Engineering and Technology. Her research interests include Cyber Security, Networking and Deep Learning. (E-mail: gkaur6_be20@thapar.edu)

Bakul Gupta is currently an undergraduate 4th year Computer Engineering student at Thapar Institute of Engineering and Technology. His research interests include Blockchain, Deep Learning, Cryptography and System Security. (E-mail: bgupta1_be20@thapar.edu)

Ashima Anand is currently working as an Assistant Professor in Thapar Institute of Engineering and Technology. She pursued her Ph.D. and MTech. in Computer Science and Engineering from NIT Patna, Bihar, India. Also, she received B. Tech. in Computer Science and Engineering from NIT Hamirpur, H.P., India in 2017. Her research interest includes Data Hiding Techniques, Cryptography and Image Processing. (E-mail: ashima.anand@thapar.edu)